ACTGTCTGGAACTGGACTGAGTCACCAAAAGGCGAATGGCTTCATCTTATAAAATGTCTGAACAAAGCACA
ACTTCTGAGCACATTTTACAGAAGACATGTGATCACCTGATCCTGACTAACCGTTCTGGATTAGAGACAGA
CTCAGTAGCAGAGGAAATGAAGCAGACTGTGGAGGGACAGGGGCATACAGTGCACTGGGCAGCTCTCCTGA
TACTCGCGGTGATAATACCCACCATTGGTGGGAACATCCTTGTGATTCTGGCTGTTGCACTGGAGAAAAGG
CTGCAGTACGCTACCAACTACTTTTTAATGTCCTTGGCGATAGCAGATTTGCTGGTTGGATTGTTTGTGAT
GCCGATTGCCCTCTTGACAATCATGTTTGAGGCTATATGGCCCCTCCCACTGGCCCTGTGTCCTGCCTGGT
TATTCCTCGATGTTCTCTTTTCAACTGCCTCCATCATGCATCTCTGTGCCATTTCCCTGGACCGCTATATA
GCCATCAAAAAGCCAATTCAGGCCAATCAGTGCAACACCCGGGCTACTGCATTCATCAAGATTACAGTGGT
ATGGTTAATTTCAATAGGCATCGCCATCCCAGTCCCTATTAAAGGAATCGAGACTGATGTGATTAATCCAC
ACAATGTCACCTGTGAGCTGACAAAGGACCGCTTTGGCAGTTTTATGGTCTTTGGGTCACTGGCTGCTTTC
TTCGTACCTCTCACCATCATGGTAGTCACTTACTTTCTCACCATTCACACTTTACAGAAGAAAGCTTACTT
GGTCAAAAATAAGCCACCTCAACGCCTAACACGGTGGACTGTGCCCACAGTTTTCCTAAGGGAAGACTCAT
CCTTTTCATCACCAGAAAAGGTGGCAATGCTGGATGGGTCTCACAGGGATAAAATTCTACCTAACTCAAGT
GATGAGACACTTATGCGAAGAATGTCCTCAGTTGGAAAAAGATCAGCCCAAACCATTTCTAATGAGCAGAG
AGCCTCGAAGGCCCTTGGAGTCGTGTTTTTCCTTTTTCTGCTTATGTGGTGCCCCTTTTTTATTACAAATC
TAACTTTAGCTCTGTGTGATTCCTGCAATCAGACCACTCTCAAAACACTCCTGGAGATATTTGTGTGGATA
GGCTACGTTTCCTCGGGGGTGAATCCTCTGATCTATACACTCTTCAATAAGACATTTCGGGAAGCATTTGG
CAGGTACATCACCTGCAATTACCGAGCCACAAAGTCAGTAAAAGCACTTAGGAAGTTTTCCAGTACACTTT
GTTTTGGGAATTCAATGGTAGAAAACTCTAAATTTTTCACAAAACATGGAATTCGAAATGGGATCAACCCT
GCCATGTACCAGAGCCCAATGAGGCTCCGATGTTCAACCATTCAGTCCTCATCAATCATCCTCCTCGATAC
CCTTCTCACTGAAAACGATGGCGACAAAGCGGAAGAGCAGGTCAGCTACATATTGCAGGAACGGGCCGGCC
TCATCTTGAGAGAGGGTGATGAGCAGGACGCACGCGCACCATGGCAGGTTCAAGAGTGA (SEQ ID NO:1)

MASSYKMSEQSTTSEHILQKTCDHLILTNRSGLETDSVAEEMKQTVEGQGHTVHWAALLILAVIIPTIGGN
ILVILAVALEKRLQYATNYFLMSLAIADLLVGLFVMPIALLTIMFEAIWPLPLALCPAWLFLDVLFSTASI
MHLCAISLDRYIAIKKPIQANQCNTRATAFIKITVVWLISIGIAIPVPIKGIETDVINPHNVTCELTKDRF
GSFMVFGSLAAFFVPLTIMVVTYFLTIHTLQKKAYLVKNKPPQRLTRWTVPTVFLREDSSFSSPEKVAMLD
GSHRDKILPNSSDETLMRRMSSVGKRSAQTISNEQRASKALGVVFFLFLLMWCPFFITNLTLALCDSCNQT
TLKTLLEIFVWIGYVSSGVNPLIYTLFNKTFREAFGRYITCNYRATKSVKALRKFSSTLCFGNSMVENSKF
FTKHGIRNGINPAMYQSPMRLRCSTIQSSSIILLDTLLTENDGDKAEEQVSYILQERAGLILREGDEQDAR
APWQVQE (SEQ ID NO:2)

FIGURE 1

Underlined = deleted in targeting construct

Bold = sequence flanking Neo insert in targeting construct

ACTGTCTGGAACTGGACTGAGTCACCAAAAGGCGAATGGCTTCATCTTATAAAATGTCTG
AACAAAGCACAACTTCTGAGCACATTTTACAGAAGACATGTGATCACCTGATCCTGACTA
ACCGTTCTG<u>GATTAGAGACAGACTCAGTAGCAGAGGAAATGAAGCAGACTGTGGAGGGAC</u>
<u>AGGGGCATACAGTGCACTGGGCAGCTCTCCTGATACTCGCGGTGATAATACCCACCATTG</u>
<u>GTGGGAACATCCTTGTGATTCTGGCTGTTGCACTGGAGAAAAGGCTGCAGTACGCTACCA</u>
<u>ACTACTTTTTAATGTCCTT</u> **GGCGATAGCAGATTTGCTGGTTGGATTGTTTGTGATGCCGA**
TTGCCCTCTTGACAATCATGTTTGAGGCTATATGGCCCCTCCCACTGGCCCTGTGTCCTG
CCTGGTTATTCCTCGATGTTCTCTTTTCAACTGCCTCCATCATGCATCTCTGTGCCATTT
CCCTGGACCGCTATATAGCCATCAAAAAGCCAATTCAGGCCAATCAGTGCAACACCCGGG
CTACTGCATTCATCAAGATTACAGTGGTATGGTTAATTTCAATAGGCATCGCCATCCCAG
TCCCTATTAAAGGAATCGAGACTGATGTGATTAATCCACACAATGTCACCTGTGAGCTGA
CAAAGGACCGCTTTGGCAGTTTTATGGTCTTTGGGTCACTGGCTGCTTTCTTCGTACCTC
TCACCATCATGGTAGTCACTTACTTTCTCACCATTCACACTTTACAGAAGAAAGCTTACT
TGGTCAAAAATAAGCCACCTCAACGCCTAACACGGTGGACTGTGCCCACAGTTTTCCTAA
GGGAAGACTCATCCTTTTCATCACCAGAAAAGGTGGCAATGCTGGATGGGTCTCACAGGG
ATAAAATTCTACCTAACTCAAGTGATGAGACACTTATGCGAAGAATGTCCTCAGTTGGAA
AAAGATCAGCCCAAACCATTTCTAATGAGCAGAGAGCCTCGAAGGCCCTTGGAGTCGTGT
TTTTCCTTTTTCTGCTTATGTGGTGCCCCTTTTTTATTACAAATCTAACTTTAGCTCTGT
GTGATTCCTGCAATCAGACCACTCTCAAAACACTCCTGGAGATATTTGTGTGGATAGGCT
ACGTTTCCTCGGGGGTGAATCCTCTGATCTATACACTCTTCAATAAGACATTTCGGGAAG
CATTTGGCAGGTACATCACCTGCAATTACCGAGCCACAAAGTCAGTAAAAGCACTTAGGA
AGTTTTCCAGTACACTTTGTTTTGGGAATTCAATGGTAGAAAACTCTAAATTTTTCACAA
AACATGGAATTCGAAATGGGATCAACCCTGCCATGTACCAGAGCCCAATGAGGCTCCGAT
GTTCAACCATTCAGTCCTCATCAATCATCCTCCTCGATACCCTTCTCACTGAAAACGATG
GCGACAAAGCGGAAGAGCAGGTCAGCTACATATTGCAGGAACGGGCCGGCCTCATCTTGA
GAGAGGGTGATGAGCAGGACGCACGCGCACCATGGCAGGTTCAAGAGTGA

FIGURE 2A

Gene Sequence
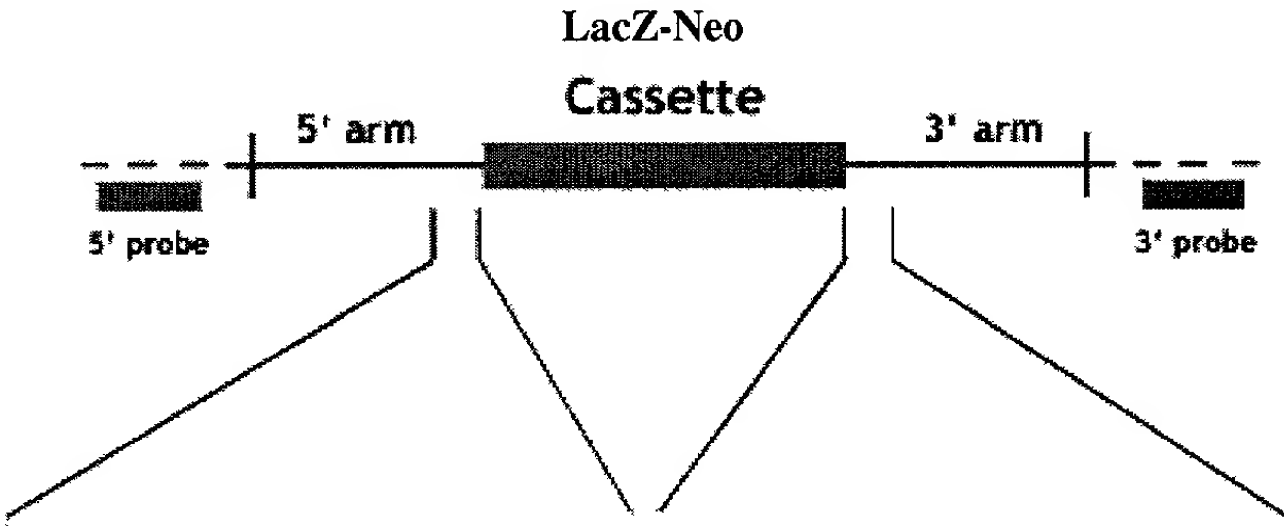Structure *

130 bp   Sequence Deleted   319 bp

Size of full-length
cDNA: 1550 bp

Targeting Vector* (genomic sequence)

Construct Number: 2520

LacZ-Neo
Cassette

5' arm   3' arm

5' probe   3' probe

Arm Length:
5': 1.6 kb
3': 5 kb

──────── Targeting Vector
– – – – – Endogenous Locus

* Not drawn to scale

5'>TGAGTGTCTGGTGGGTTTGCT
AAATGCTTTGCTAAAGCAGATGAC
TTGCTTAGCTACTGACCATGCTGA
CCACTGTCTGGAACTGGACTGAGT
CACCAAAAGGCGAATGGCTTCATC
TTATAAAATGTCTGAACAAAGCAC
AACTTCTGAGCACATTTTACAGAA
GACATGTGATCACCTGATCCTGAC
TAACCGTTCTG<3'
(SEQ ID NO:3)

5'>GGCGATAGCAGATTTGCTGGT
TGGATTGTTTGTGATGCCGATTGC
CCTCTTGACAATCATGTTTGGTGA
GTATTTCCCCTTGTTCCTGCCACT
GAACACTACTAACGTAGTGAAATG
GACACTCACTGACCTTTATTTTGT
TTGAAATAAAAGAAGGACCTGGAT
TAAAAACACAGAAGGGAACATTCC
TTCATTTTTCA<3'
(SEQ ID NO:4)

FIGURE 2B